National Tax Journal, September 2015, 68 (3S), 875–900 http://dx.doi.org/10.17310/ntj.2015.3S.07

SCHOOL ACCOUNTABILITY AND SCHOOL CHOICE: EFFECTS ON STUDENT SELECTION ACROSS SCHOOLS

Cassandra M. D. Hart and David N. Figlio

This paper examines the effect of the introduction of school accountability policies in Florida on schools' student body composition. We specifically examine the effects of the state issuing official school "grades" on the composition of incoming kindergarten classes, using novel data on families' socioeconomic characteristics drawn from birth records. High socioeconomic status parents were particularly responsive to the introduction of grades. Schools that received A grades saw significant increases in an index measure of socioeconomic status among their kindergarten students after enactment of the policy. We find some evidence that responses are stronger for A schools that have nearby alternatives, and where nearby alternatives are poorer-performing schools.

Keywords: school accountability, school choice, information shocks

JEL Codes: I24, I28, H730

I. INTRODUCTION

Researchers have recognized peer effects as an important influence on student achievement and academic attainment since the issuance of the Coleman Report (Coleman, 1968). Research in recent years has established causal relationships between peer qualities and student outcomes, with higher-achieving peers (Hoxby and Weingarth, 2005; Lavy, Paserman, and Schlosser, 2012; Imberman, Kugler, and Sacerdote, 2012) and less disruptive peers (Figlio, 2007; Carrell and Hoekstra, 2010) associated with higher achievement.[1] The positive effect of high-achieving peers seems to be especially strong for the lowest-achieving students (Burke and Sass, 2013; Sojourner, 2013).

Because parental socioeconomic characteristics are strongly and increasingly associated with student achievement (Reardon, 2011), policymakers are often concerned

[1] Epple and Romano (2011) provide a comprehensive overview of the literature on peer effects in education.

Cassandra M. D. Hart: School of Education, University of California, Davis, Davis, CA, USA (cmdhart@ucdavis.edu)

David N. Figlio: Institute for Policy Research, Northwestern University, Evanston, IL, USA, and National Bureau of Economic Research, Cambridge, MA, USA (figlio@northwestern.edu)

about policies that may increase stratification by socioeconomic status across schools. In particular, effects on peer composition and concerns of "cream-skimming" have received a lot of attention in the school choice literature, and have been studied with respect to the use of both public (Bifulco, Ladd, and Ross, 2009; Zimmer et al., 2009; Cowen and Winters, 2013) and private (Figlio, Hart, and Metzger, 2010; Chakrabarti, 2013) school choice options. However, there has been little work looking at the effects on peer composition of one of the most influential policy initiatives of the last two decades: school accountability policies. This paper studies whether the public reporting of school grades under the school accountability policy in Florida is associated with changes in the overall enrollment and composition of students across schools that receive different grades.[2] Specifically, we investigate whether the repackaging of already-available information by the state into accountability scores for schools has an independent effect on student sorting, above and beyond the sorting associated with any of the underlying factors that contribute to the school grades. In doing so, we follow the approach introduced by Figlio and Lucas (2004) to study the effects of the introduction of Florida's school accountability system on the housing market.

Accountability systems may affect the sorting of students between schools in one of three ways. First, it is possible that the publication of school grades will cause the most response among low socioeconomic status (SES) parents. For instance, in the absence of state-provided information on school quality, more informed, educated parents may have an advantage in determining school quality. This would be true if they had better access to networks that provide informal information on school quality. In this case, state provision of a new, low-cost source of information on school quality could level the playing field, and there may be more reaction to the introduction of new information among lower-income, less educated families as they gain information that had already been incorporated into the education decisions of higher-SES families. A second alternative is that high-SES families may respond most strongly to the new information. If high-SES families simply have an advantage in acting on school quality information — for instance, because they are more likely to have the financial latitude to move closer to higher-ranked schools — then provision of additional information might deepen existing inequalities and result in more socioeconomic segregation across schools. A third possibility is that parents will not respond strongly to the new information at all. This would be likely either if parents found the new information difficult to interpret, if parents were uninterested in academic quality, or if they were already fully aware of the new information.

A healthy body of literature has examined how parents make decisions about where to send their children in the context of school choice. In particular, parents report in interviews that they place high value on academic quality when determining whether to participate in school choice programs (Witte, 2001; Greene, Howell, and Peterson, 1997; Beales and Wahl, 1995). However, empirical studies that attempt to tease out

[2] For reviews on the performance and behavioral consequences of school accountability policies, see Figlio and Ladd (2008) and Figlio and Loeb (2011).

the importance of academic quality to parents have come to mixed conclusions about whether parents' behavior is consistent with these reports. Examining patterns of parent searches on information about school quality, Schneider and Buckley (2002) find that parents tend to search for information on peer composition ahead of academic quality, although higher-SES parents' search patterns suggested more interest in academic quality. Using several sources of data, Rothstein (2006) finds little evidence that higher-SES families cluster near more effective schools. However, Rothstein's study examined a pre-accountability period (for most states), and he acknowledges that the increased provision of information under state accountability systems could affect the extent to which parental sorting around high-quality schools occurred. Indeed, experimental evidence suggests that direct provision of information on school quality to parents, especially when the presentation is designed to increase the salience of the information, can affect parents' schooling decisions for their children; Hastings and Weinstein (2008) find that when low-income parents are provided with information on the test scores of surrounding schools, they are more likely to choose higher-scoring alternatives for their children.

The study closest to ours provides a look at compositional effects of the introduction of school grades on a sample of cross-metropolitan movers between one set of unidentified school districts (Figlio and Lucas, 2004). That study found that schools newly identified as A schools attracted students with higher prior reading scores, but no measurable difference in the likelihood of subsidized lunch eligibility, in the year after the introduction of the school grades than they had in the year prior to the policy change; however, by the next year, the effect of A receipt on the new student body was more tempered. There is also evidence that school accountability systems may influence neighborhood composition in different ways when school choice is linked to accountability: Billings, Brunner, and Ross (2014) find that higher-income families moved to school zones of "failing" schools in North Carolina once residence in these neighborhoods gave students an advantage in lotteries to attend high-demand schools. While this latter result does not have direct implications for school compositional changes as a result of accountability, it does provide further evidence that households are highly responsive to the information conveyed by accountability systems.

Our study extends these results in several important ways. First, we are able to examine a richer set of students' family characteristics than other studies because we have a unique data set of birth records. The measure of socio-economic status available in administrative education data, eligibility for free or reduced-price lunch, is a weak proxy for family status because it combines the working poor with those in extreme poverty, and the near-poor with very affluent families. In our case, we can study differential selection by variables such as parental education. Furthermore, our analysis implicitly looks at changes in compositions that come from parent decisions to alter school attendance within district, as well as differences in schools' ability to attract students carrying out inter-district moves. In addition, we examine heterogeneity in compositional effects based on differences in the availability and quality of other nearby schools.

A. School Accountability in Florida

Florida introduced its state-level accountability program, the A+ Accountability Plan, in spring 1999. Prior to the A+ Plan's passage, parents were provided with limited information about school performance. For instance, some newspaper reports of Florida Comprehensive Assessment Test (FCAT) scores in spring 1998 gave information including mean scores or proficiency rates of local schools relative to district and state averages, but the explanations of what these numbers meant were often unclear or relatively complex for lay audiences.[3] Signals about very poor performing schools were relatively clearer, with the state establishing a list of critically low-performing schools (Figlio and Lucas, 2004). However, even in that case, the critically low-performing schools list was based on discretionary factors at the school district level.

The A+ Plan was passed in January 1999 to increase school accountability for student performance. The plan established a formula to determine schools' letter grades on an A–F scale based on aggregate student performance (Florida Department of Education, 1999). Students were categorized into one of five levels on each subject depending on their test scores. Level 2 denoted limited success in a given subject, while level 3 denoted partial success (i.e., challenge meeting the most difficult grade-level standards) (Consortium for Policy Research in Education, 2000).

Grades from C–F were determined based on how a school's students performed on a set of "minimum" criteria on each of three subjects: reading, writing, and math. The minimum criteria for both math and reading required 60 percent of students to score at level 2 or above; the minimum criterion for writing required 50 percent of students to score at level 3 or above. Schools that failed on all three criteria received F grades; schools that met at least one, but not all, criteria, received D grades, and schools that met all of these standards, but not the higher-level criteria, received C grades.

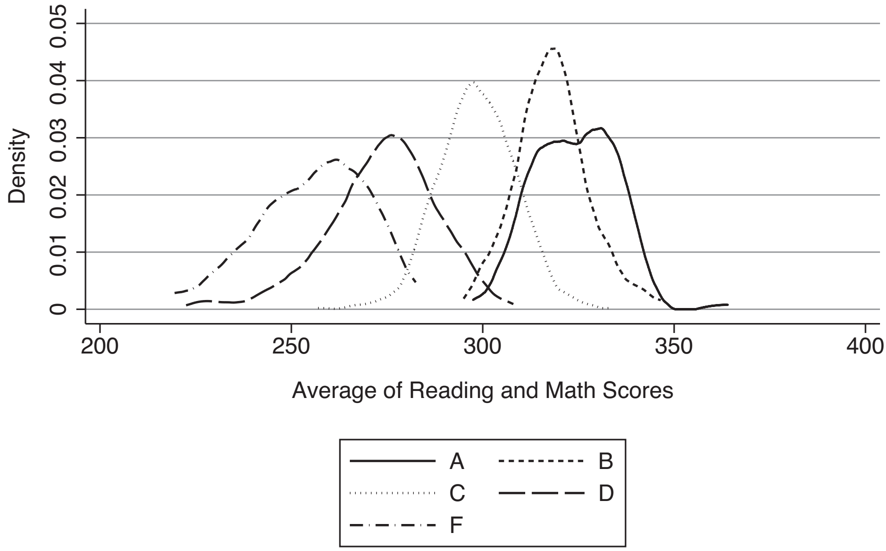
Schools had to meet a higher set of standards to receive a B or an A. To meet the higher-performing thresholds in reading or math, 50 percent of students had to score at level 3 or above. Sixty-seven percent of students had to score at level 3 or better to meet the higher-performing writing criterion. To receive a B, schools had to meet higher-performing thresholds in all subjects; no subgroup could fall below minimum criteria, and testing rates had to exceed 90 percent for standard-curriculum students. To receive an A, schools additionally had to post absentee rates and suspension rates lower than the state average; achieve a substantial improvement in reading compared to the prior-year test scores; *not* decline in writing and math; and test at least 95 percent of standard curriculum students.

Figure 1 plots the density of the school average FCAT scores for schools that received each grade A–F. The average FCAT score for each school represents the mean of the school's average math and reading scores. Given the complexity of the testing criteria

[3] For instance, a St. Petersburg newspaper listed local schools' scores for 4th, 5th, 8th, and 10th grade math and reading tests, and the number of points above or below the grade-subject specific state average, but provided no context on what these point differences meant (Chion-Kenney, 1998).

Figure 1

Density of Averaged Math and Reading Scores by 1999 Grade
Elementary Schools 1998–1999



Note: Standard Deviation=24.69

at the higher level in particular, it was easy for schools to narrowly receive B ratings despite having observationally similar testing performance to A schools; this is evident in the substantial overlap of the distributions of the B and A schools in Figure 1.

II. DATA, MEASURES, AND METHODS

A. Data

We analyze characteristics of incoming kindergarten classes from the full population of 1,412 elementary schools in Florida using several different data sources. Data on school enrollment for individual students are drawn from individual student records maintained by the Florida Department of Education. These student records are merged to an unusually rich set of data on family characteristics drawn from birth records for students entering kindergarten from fall 1997 through fall 2001. The birth records were provided by the Florida Department of Health for all children born between 1992 and 2002. Birth and school records were merged on name, date of birth, and Social Security

Number. In total, 80.7 percent of births were matched to public school records; the range that would have been expected based on American Community Survey data is between 79–81 percent, depending on assumptions about how many people leave the country before age 5.[4] These merged birth-and-school records, representing 528,516 kindergarten enrollments,[5] were used to generate school-level measures reflecting the characteristics of elementary schools' incoming kindergarten classes. Because we rely on birth records to provide information on many of the characteristics we study, the results are calculated using only students that we have birth records for, i.e., Florida natives.[6]

B. Models

We use a difference-in-differences design to examine the effect of schools' receipt of different levels of grades in 1999 (A, B, C, D, or F) on incoming kindergarten class characteristics.[7] Kindergarten classes are used because parents may be less responsive to new information when their children are in later grades and have established relationships in a school.

To illustrate, to determine the effect of receiving an A (versus all other grades) in 1999 on kindergarten class composition, we estimate the following equation for incoming kindergarten classes from the 1997–1998 school-year through the 2001–2002 school-year:

$$(1) \quad K_{sdt} = \beta \boldsymbol{Year} \times A_{sd,1999} + \rho \boldsymbol{Year}_d + \delta \boldsymbol{SchoolVars}_{sd,1999} + \tau_s + \varepsilon_{st},$$

where K is a kindergarten class composition measure in year t for school s in school district d, $Year \times A$ represents a series of interaction effects for different years, interacted with indicators for whether schools received an A grade in 1999, *Year* is a vector of school district-specific year dummies, *SchoolVars* is a vector of controls for all variables used to determine school grades, measured at the 1999 baseline when school grades were first constructed, and τ is a school fixed effect (which implicitly absorbs the 1999 levels of the school variables).[8] We fix our attention on the 1999 grade as an event study

[4] See Figlio et al. (2014) for details regarding the nature of these data and the match between school and birth records in Florida.

[5] This is the number of student enrollments in kindergarten in our study timeframe, including multiple enrollments for children who repeat kindergarten.

[6] Results for enrollment are substantively similar when we conduct a separate check using enrollment figures drawn from the Common Core of Data maintained by the National Center for Education Statistics.

[7] This is the empirical approach applied by Figlio and Lucas (2004) and Figlio and Rouse (2006) regarding the 1999 introduction of school accountability in Florida. Studies of a subsequent (2002) change in the school accountability system, including Chiang (2009), Figlio and Kenny (2009), Rouse et al. (2013), and West and Peterson (2006), employ a regression discontinuity design to study the effects of receipt of a particular grade, but the nature of the initial accountability system does not permit analysis using a regression discontinuity design because of its complicated formula for assignment of specific school grades.

[8] We have also estimated these models in first difference specifications rather than the fixed effects estimators, where the first differences are measured, successively, in differences between the baseline year and the year in question. The results of the first difference specifications are very similar to those reported herein.

because this was the first time that the state explicitly evaluated schools. We consider the 1999 event rather than year-by-year changes in school grades because, as Figlio and Lucas (2004) document, there was a substantial degree of churning in school grades in the years immediately following the introduction of the school accountability policy.[9]

Our coefficients of interest are the β's, which should be negligible in pre-policy years and significant and consistently signed afterwards if there are measurable responses to A grades. Importantly, since we control for interactions between year and the school variables that underlay the school grades in 1999, we are explicitly studying the effects of a school's receipt of an A grade, say, on differential sorting over time, above and beyond any sorting that is associated with the underlying variables themselves. We cluster our standard errors at the school level.

Each grade group is compared to schools that received grades between the referent grade and a C. That is, A (F) schools are compared to B and C (D and C) schools, while B (D) schools are compared only to C schools. There are two reasons for these restrictions. First, we want to ensure that our comparisons for B and D schools are made cleanly to schools that clearly perform worse (better) than they do; including A and F schools for these groups respectively muddies the counterfactual. Second, we exclude schools that are at the opposite end of the grade distribution because of the relative lack of overlap in the underlying test scores of A/B versus D/F schools. There is relatively little area of common support between schools at the very top and the very bottom of the distribution (Figure 1).[10] Because of the way our restrictions are designed, we do not estimate effects for C schools.

C. Measures

We use a wide range of dependent variables to determine if the characteristics of kindergarten classes changed in response to the provision of school grades. From the birth certificate data, we observe mothers' years of completed education, mothers' age at the child's birth, mothers' marital status at birth, and mothers' number of previous live births. We use this information to calculate the share of the kindergarten class

---

[9] In the conference version of this paper, we controlled for year-by-year lagged values of the school characteristics that went into the grade determinations, following Figlio and Lucas (2004). The general pattern of the results is similar, especially with regard to the A grade treatment, which remains our most robust set of findings. We have also estimated models that include interactions between year dummies and 1999 school variables. The pattern of results for enrollment figures are substantively similar, although the magnitudes of the effects on the socioeconomic status mix variables are often less precisely estimated. We find similar results when we instead use the school variables for 1998 and control instead for interactions between year dummies and 1998 school variables. We prefer the models that we present here because we believe that the presentation of grades might heighten the salience of the test information in the post-policy years, resulting in over-controlling for the *Grade* × *Year* effects.

[10] We have also run models that do not throw out the schools at the opposite end of the distribution, models that compare A, B, D, and F schools only to C schools, and models that compare schools only to nearest-neighbor grades towards the middle of the distribution (e.g., A to B, B to C, D to C, F to D). Under all of these specifications, the main findings are substantively similar.

with married parents at birth, the average maternal age at birth, and the average years of completed education for mothers at birth. Information on students' free and reduced lunch status was obtained from the Florida Department of Education (FDOE). From that, we generate information on the share of incoming classes that are low-income, i.e., that use subsidized lunch. Finally, we characterize enrollment figures based on the (log) number of total students entering each kindergarten class.[11] Data on school grades assigned by the Department of Education were obtained from the Department of Education through the Florida School Accountability Reports (Florida Department of Education, 2013). Controls, including school average test scores on reading and math, were provided by the Department of Education, as well as information on absentee rates, suspension rates, and mobility/stability rates. The latter three measures were used in conjunction with test scores to determine overall school grades.

III. RESULTS

Descriptive statistics for all measures as of 1998–99 are provided in Table 1, broken down by the school grade received in spring 1999. Not surprisingly, schools that received different grades had notably different characteristics in the fall prior to the initial issuing of grades, with schools' students exhibiting lower average maternal education rates, younger maternal ages at birth, and lower likelihood of mothers having been married at birth as school grades descend from A to F. These differences across grade levels highlights the importance of using school fixed effects to ensure that effects are identified from within-school changes in composition following grade announcements.

At the same time, note that A and B schools look quite similar on most dimensions. Because the Florida accountability system was originally structured as a conjoint system in which schools had to meet all of several requirements to receive A grades, schools that looked quite similar along most characteristics could end up with different grades due to a failure to meet targets on just one measure.

A. School Composition

To explore whether school composition changed in response to the 1999 accountability information shocks, we run grade-by-year fixed effect models for each of four grade groups (A, B, D, and F) on each of five outcomes (average years of maternal education, average maternal age at child's birth, share of mothers married at child's birth, share of students on subsidized lunch, and share of students whose mothers were not born in the United States). Interactions for *Grade × 1998* are omitted; therefore, all coefficients can be interpreted as the differential change between schools that received the relevant grade and comparison schools between the referent year and 1998. The

[11] Measures are available for the estimated three-quarters of the cohort starting kindergarten in 1997 that were born in 1992; entering kindergartners born in September-December 1991 or red-shirted kindergartners are not captured by our measures. We address this point later in the paper.

Table 1

Descriptive Statistics as of Fall 1998, by 1999 School Grade

	A Mean (Standard Deviation)	B Mean (Standard Deviation)	C Mean (Standard Deviation)	D Mean (Standard Deviation)	F Mean (Standard Deviation)
Characteristics of Kindergarten Class: School Means					
Maternal years education	13.450	13.175	12.210	11.293	11.110
	(0.778)	(0.822)	(0.798)	(0.936)	(0.770)
Maternal age	28.459	27.934	26.271	25.204	24.655
	(1.587)	(1.655)	(1.547)	(1.509)	(1.335)
Mother married (%)	81.907	78.798	66.296	45.904	32.589
	(9.765)	(10.307)	(12.879)	(15.684)	(14.764)
Using FRL[1] (%)	21.924	26.365	46.971	73.311	85.329
	(14.219)	(15.945)	(19.326)	(16.827)	(13.163)
Mother non-US-born (%)	12.846	14.985	18.898	29.770	27.904
	(11.990)	(11.827)	(20.245)	(25.941)	(24.510)
Number of entering K students	88.219	87.220	92.651	98.261	89.098
	(24.151)	(27.020)	(30.923)	(39.604)	(30.545)
School Performance Measures					
School average reading score	317.88	309.21	291.81	263.15	243.47
	(12.344)	(10.232)	(12.134)	(16.674)	(16.628)
School average math score	330.27	326.36	305.60	281.79	264.28
	(9.799)	(9.561)	(10.478)	(15.805)	(14.520)
Share absent 21-plus days	4.215	5.275	7.739	8.961	9.524
	(1.736)	(2.252)	(2.759)	(3.208)	(3.669)
Share students with ISS[1]	0.654	0.837	1.491	1.771	3.056
	(1.663)	(1.846)	(3.008)	(3.628)	(6.422)
Share students with OSS[1]	0.824	1.072	1.810	3.340	4.650
	(1.018)	(1.218)	(1.872)	(3.282)	(3.660)
Mobility rate	21.383	23.558	33.697	48.828	55.603
	(9.143)	(10.208)	(11.614)	(42.474)	(19.920)
Unique schools	114	192	639	397	54

Notes: Spring 1998 characteristics are based on 1999 school grades and weighted by 1998 number of students. Maternal characteristics are measured at birth. Free/reduced lunch (FRL) status is measured in kindergarten year.

[1]FRL = Free/Reduced Lunch, ISS=In-school suspensions, OSS=Out-of-school suspensions

pattern that we would expect to see if receiving a given grade prompts a change in the school composition feature captured by the dependent variable is a null result for the *Grade × 1997* coefficient, along with appropriately signed, significant coefficients for the *Grade × 2000,* and *Grade × 2001* variables. Because grades were released in June, leaving parents with relatively little time to react to school grades by changing their child's enrollment, we are somewhat more agnostic about the likelihood that the *Grade × 1999* variables should be significant. Likewise, because new grades come out each year, we acknowledge that by summer 2001, when parents were making their enrollment decisions for next fall, they had multiple years of grades to weigh. The 1999 grades may have been less salient in this year as parents attended to performance in subsequent years, and school grades had the tendency to vary from year to year, so we may expect some attenuation of the effect of 1999 A grades by this year.

Our first results are presented in Table 2. Each cell represents the interaction coefficient for the grade referenced in the column heading with the year referenced in the row heading. The comparison group is also denoted in the column heads. Panel A suggests that the average maternal education of a school's entering kindergarten class changes in response to the receipt of an A grade. Specifically, we see significantly greater increases in the average maternal education of schools in 1999, 2000, and 2001 compared to the 1998 levels for A schools than for comparison (B/C) schools (Column 1). We observe more modest but positive estimated effects of a B grade (relative to a C grade) on maternal education as well (Column 2). By comparison, we observe reductions in maternal education associated with an F grade in 1999, relative to a C or D grade (Column 4). If anything, the estimated effects of the initial grade rise modestly over time.

The patterns of coefficients for the A and B comparisons are similar with regard to maternal age (Panel B of Table 2), though not for the F comparison. There is, however, less evidence that receipt of any grade in particular affects the composition of the kindergarten class in terms of the share of mothers who are married (Table 2, Panel C) or the share of students who are using free or reduced price lunch (Table 2, Panel D). There are scattered cases of significant coefficients, but the patterns are not sufficiently strong to conclude that the announcement of grades systematically altered the composition of schools on these dimensions. The patterns of results regarding the fraction of foreign-born mothers (Table 2, Panel E) are roughly comparable to those regarding maternal education and age, though less frequently statistically distinct from zero.

Presenting results for all five outcomes throughout the paper would be expositionally onerous, and each of these measures is itself only a weak proxy for SES, so we used third grade tests scores for students to empirically generate an "advantage index" for each student.[12] Specifically, we standardized students' third-grade math and read-

[12] This is nearly identical to the approach employed by Figlio et al. (2014) to study the differential effects of birth weight on test scores for children from different socio-economic backgrounds.

Table 2

Effect of 1999 Grade Received on School Kindergarten Composition, Year-by-Year, School Fixed Effects

	1	2	3	4
	A [Comparison Group: B, C]	B [Comparison Group: C]	D [Comparison Group: C]	F [Comparison Group: C, D]
Panel A: Mean Maternal Education				
Grade × 1997	–0.011	–0.038	0.056**	–0.040
	(0.042)	(0.035)	(0.027)	(0.053)
Grade × 1999	0.073**	0.030	–0.020	–0.096*
	(0.032)	(0.029)	(0.026)	(0.050)
Grade × 2000	0.129***	0.067*	–0.079***	–0.139*
	(0.035)	(0.034)	(0.026)	(0.078)
Grade × 2001	0.153***	0.072**	–0.038	–0.163**
	(0.043)	(0.034)	(0.029)	(0.083)
Group outcome mean	13.450	13.189	11.308	11.110
Group outcome standard deviation	0.778	0.814	0.903	0.770
Panel B: Mean Maternal Age				
Grade × 1997	–0.045	0.047	0.028	0.145
	(0.095)	(0.088)	(0.069)	(0.127)
Grade × 1999	0.134	0.091	–0.080	0.005
	(0.085)	(0.078)	(0.060)	(0.129)
Grade × 2000	0.233**	0.145	–0.150**	0.116
	(0.090)	(0.091)	(0.064)	(0.146)
Grade × 2001	0.295***	0.266***	–0.106	–0.123
	(0.104)	(0.088)	(0.069)	(0.171)
Group outcome mean	28.459	27.973	25.209	24.655
Group outcome standard deviation	1.587	1.615	1.511	1.335

Table 2, Continued

	1 A [Comparison Group: B, C]	2 B [Comparison Group: C]	3 D [Comparison Group: C]	4 F [Comparison Group: C, D]
Panel C: Fraction with Married Parents				
Grade × 1997	–0.007	–0.002	–0.002	0.002
	(0.007)	(0.007)	(0.005)	(0.011)
Grade × 1999	0.008	–0.000	–0.002	0.008
	(0.005)	(0.006)	(0.005)	(0.011)
Grade × 2000	0.018***	0.009	–0.003	0.013
	(0.006)	(0.006)	(0.005)	(0.012)
Grade × 2001	0.012	0.009	0.003	0.023*
	(0.008)	(0.007)	(0.005)	(0.012)
Group outcome mean	0.819	0.788	0.459	0.326
Group outcome standard deviation	0.098	0.104	0.157	0.148
Panel D: Fraction on Subsidized Lunch				
Grade × 1997	0.001	0.006	–0.008	–0.012
	(0.009)	(0.007)	(0.007)	(0.014)
Grade × 1999	–0.007	–0.006	–0.003	–0.009
	(0.007)	(0.006)	(0.006)	(0.014)
Grade × 2000	–0.033***	–0.003	0.009	–0.007
	(0.008)	(0.008)	(0.007)	(0.018)
Grade × 2001	–0.014	–0.011	–0.002	–0.017
	(0.009)	(0.008)	(0.008)	(0.018)
Group outcome mean	0.219	0.261	0.732	0.853
Group outcome standard deviation	0.142	0.158	0.168	0.132

Table 2, Continued

	1	2	3	4
	A [Comparison Group: B, C]	B [Comparison Group: C]	D [Comparison Group: C]	F [Comparison Group: C, D]
Panel E: Fraction with non-US native Mothers				
Grade × 1997	0.010	0.003	0.002	–0.003
	(0.006)	(0.006)	(0.004)	(0.011)
Grade × 1999	–0.009*	–0.006	0.006	0.004
	(0.005)	(0.005)	(0.004)	(0.008)
Grade × 2000	–0.009*	–0.011**	0.001	0.011
	(0.005)	(0.006)	(0.004)	(0.011)
Grade × 2001	–0.008	–0.017***	0.000	0.005
	(0.006)	(0.005)	(0.005)	(0.010)
Group outcome mean	0.128	0.151	0.297	0.279
Group outcome standard deviation	0.120	0.118	0.259	0.245
School controls	Yes	Yes	Yes	Yes
School fixed effects	Yes	Yes	Yes	Yes
District-year fixed effects	Yes	Yes	Yes	Yes
Unique schools	943	829	1,033	1,087
School-years	4,693	4,123	5,145	5,415

Notes: Asterisks denote significance at the 1% (***), 5% (**), and 10% (*) levels. Data are weighted by size of kindergarten entering class, fall 1998. Years represent fall of academic year. Year 1998 was omitted. Grades were introduced June 1999. Robust standard errors are clustered at school level. Group outcome mean and standard deviation are calculated for group that received the referent grade labeled in column header, in fall of 1998 (weighted by fall 1998 kindergarten student counts).

ing FCAT scores within year for a sample of students observed in third grade from roughly 2000–2010.[13] Third grade is the earliest year that state standardized tests are administered in Florida. We then regress a variable that averages the math and reading standardized FCAT scores on a series of birth or other characteristics that are exogenous to third-grade test performance — maternal education, marital status, immigrant status, and age at birth; student race and kindergarten subsidized lunch use status as recorded

[13] This represents the full sample of students for whom we have both birth record data and public school data.

by the district.[14] We then predict third grade average FCAT standardized scores using these exogenous characteristics for each student. These predicted FCAT values can be thought of as an index of advantage in that they capture the FCAT scores that would be predicted purely based on students' socioeconomic characteristics. We use this index variable as the main outcome of interest for the remainder of the paper.[15] The mean index value for each school is strongly correlated ($p < 0.001$) with each of the mean component variables in bivariate correlations.

Table 3, Panel A presents the results when this index is used as the outcome variable. Results are largely similar to when maternal education is used as the outcome variable in Table 2; we find little persuasive evidence that receipt of B, D, or F grades is associated with changes in overall composition of schools on our advantage index. Receipt of an A, however, is associated with a significantly positive differential trend in the advantage index level of kindergarten classes in 1999, 2000, and 2001 compared to 1998 relative to the trends in B and C schools. This differential trend is not evident in the pre-policy period. Note that while these differences are statistically significant, the magnitudes of the differences are modest. For instance, the differential increase of 0.013 in 1999 (over 1998) for A schools relative to B/C schools represents about 0.10 of the standard deviation (SD) of the index measure for A schools (SD = 0.134). The index mean and standard deviation for schools in each grade category are given in the bottom rows of Panel A.

Just as we posit that changes in enrollment behavior might be more pronounced among kindergartners because older children may have ties that parents are reluctant to disrupt, we posit that first-born children might be disproportionately likely to respond to new information because any responses would not necessitate either movement of an older child or separation of two children. The birth record data allows us to observe the number of previous live births for each mother. Using only students whose mothers reported no previous live births, we re-ran the models presented in Panel A measuring the characteristics of incoming classes of first-born kindergarteners in each school. The results (Panel B) are fairly similar for the first-born and full group samples in terms of the effects of receiving A, D, or F grades. The magnitudes of the coefficients on the $A \times Year$ terms in the post-policy years tend to be slightly larger for the first-born sample, but the differences are modest. However, there are interesting differences between the two samples for schools that receive B grades; there is a significant differential increase for B schools relative to C schools in the mean advantage index of firstborn kindergarteners in the post-policy years.

[14] Race categories include Black, White, Hispanic, Asian, or other; lunch categories include free, reduced, or non-subsidized. Missing variable dummies are used to ensure information can be used for students for whom we lack information. While we do not look at race as a main outcome in the rest of the paper, we include it here because it is clearly exogenous and is strongly predictive of third grade scores. Results are substantively similar if we run an alternate specification in which the index measure is generated without using race/ethnic variables as predictors.

[15] Because the "advantage index" dependent variable is a generated variable and therefore measured with error, we have also estimated models with bootstrapped confidence intervals (Lewis and Linzer, 2005). The results are similar to those reported in Table 3. For example, the bootstrapped 95 percent confidence interval associated with *Grade × 1999* in the first column of Table 3 is [0.004, 0.022], as compared with the [0.003, 0.023] generated with ordinary least squares (OLS) estimation.

Table 3

Effect of 1999 Grade Received on Advantage Index
(Predicted Third Grade Standardized Scores), Year-by-Year, School Fixed Effects

	1 A [Comparison Group: B, C]	2 B [Comparison Group: C]	3 D [Comparison Group: C]	4 F [Comparison Group: C, D]
Panel A: Full Sample				
Grade × 1997	–0.004	–0.006	0.007*	0.003
	(0.006)	(0.005)	(0.004)	(0.009)
Grade × 1999	0.013**	0.004	–0.002	–0.001
	(0.005)	(0.005)	(0.004)	(0.008)
Grade × 2000	0.025***	0.009	–0.006	–0.001
	(0.006)	(0.006)	(0.004)	(0.012)
Grade × 2001	0.022***	0.012**	–0.003	0.012
	(0.007)	(0.006)	(0.005)	(0.011)
Group outcome mean	0.280	0.239	–0.251	–0.398
Group outcome standard deviation	0.134	0.128	0.202	0.157
Panel B: Firstborn Sample				
Grade × 1997	–0.005	0.002	0.017**	–0.006
	(0.009)	(0.009)	(0.007)	(0.014)
Grade × 1999	0.018**	0.017**	0.002	0.011
	(0.008)	(0.007)	(0.006)	(0.011)
Grade × 2000	0.029***	0.021***	–0.009	0.007
	(0.009)	(0.008)	(0.007)	(0.020)
Grade × 2001	0.031***	0.029***	–0.001	0.012
	(0.010)	(0.008)	(0.007)	(0.017)
Group outcome mean	0.290	0.246	–0.236	–0.403
Group outcome standard deviation	0.143	0.139	0.219	0.192
School controls	Yes	Yes	Yes	Yes
School fixed effects	Yes	Yes	Yes	Yes
District-year fixed effects	Yes	Yes	Yes	Yes
Unique schools	943	829	1,034	1,088
School-years	4,688	4,118	5,142	5,412

Notes: Asterisks denote significance at the 1% (***), 5% (**), and 10% (*) levels. Data are weighted by size of kindergarten entering class, fall 1998. Robust standard errors are clustered at school level. Group outcome means and standard deviations are calculated for group that received grade labeled in column header, in fall of 1998 (weighted by fall 1998 kindergarten student counts).

B. Enrollment

If parents take school grades seriously in making their children's kindergarten enrollment decisions, we would expect to observe a jump in enrollment in higher-graded schools following the grade announcements in 1999. We therefore re-run our models using logged kindergarten enrollment as the outcome variable (Table 4, Panel A). The

Table 4

Effects on Log Enrollment of 1999 Grade Received, Year-by-Year, School Fixed Effects

	1 A [Comparison Group: B, C]	2 B [Comparison Group: C]	3 D [Comparison Group: C]	4 F [Comparison Group: C, D]
Panel A: Total Enrollment				
Grade × 1997	–0.032 (0.027)	0.004 (0.020)	0.007 (0.016)	0.039 (0.028)
Grade × 1999	0.063*** (0.019)	0.040** (0.016)	0.013 (0.013)	0.004 (0.032)
Grade × 2000	0.065*** (0.025)	0.049*** (0.019)	–0.056*** (0.016)	–0.082** (0.040)
Grade × 2001	0.039 (0.026)	0.065*** (0.020)	–0.086*** (0.018)	–0.047 (0.039)
Panel B: Firstborn Enrollment				
Grade × 1997	–0.015 (0.036)	–0.021 (0.030)	–0.011 (0.024)	0.021 (0.048)
Grade × 1999	0.079*** (0.028)	0.028 (0.025)	0.014 (0.019)	–0.009 (0.043)
Grade × 2000	0.093*** (0.032)	0.007 (0.030)	–0.036* (0.022)	–0.110** (0.056)
Grade × 2001	0.058* (0.033)	0.031 (0.028)	–0.078*** (0.024)	0.003 (0.052)
School controls	Yes	Yes	Yes	Yes
School fixed effects	Yes	Yes	Yes	Yes
District-year fixed effects	Yes	Yes	Yes	Yes
Unique schools	943	829	1,034	1,088
School-years	4,688	4,118	5,142	5,412

Notes: Asterisks denote significance at the 1% (***), 5% (**), and 10% (*) levels. Data are weighted by size of kindergarten entering class, fall 1998. Robust standard errors are clustered at school level.

differential change in enrollment relative to 1998 in the post-policy years ranges from 4 to 7 percent increases for A schools relative to B/C schools, with no evidence of differences in pre-policy trends (Column 1). B schools likewise see enrollment jumps relative to C schools in the post-policy years (Column 2). By comparison, D schools appear to suffer relatively greater enrollment declines in the post-policy years relative to C schools (Column 3), beginning in 2000. The patterns are similar, though less precisely estimated, in the case of F schools relative to C/D schools (Column 4).

We check these results on a number of different samples that we think might be especially responsive to the policy. First, we re-run the models restricting our dependent variable to include the enrollment only of first-born children (Table 4, Panel B). The results for A schools and D schools are substantively similar, while the respective pattern of relative enrollment gains for B schools effectively disappears.

Because our results in Table 2 suggested that A grades were associated with an increase in average years of education and (more weakly) with a decrease of the share of students on subsidized lunch, we next re-defined our dependent variable to reflect the logged enrollment patterns for children of parents in different educational and income groups. Because our results so far have been concentrated among schools that have received A grades, we focus on that group (Table 5). Relative to these groups, A schools see significant positive post-policy changes in enrollment of children of more

Table 5

Effects on Log Enrollment for 1999 A Grade, Year-by-Year, School Fixed Effects

	1 Mother: College	2 Mother: No College	3 Not Free/Reduced Lunch	4 Free/Reduced Lunch
Grade × 1997	–0.051 (0.037)	–0.056 (0.039)	–0.032 (0.032)	–0.084 (0.056)
Grade × 1999	0.086*** (0.026)	0.031 (0.025)	0.075*** (0.022)	0.028 (0.039)
Grade × 2000	0.080*** (0.031)	–0.012 (0.033)	0.117*** (0.031)	–0.142** (0.058)
Grade × 2001	0.061* (0.031)	–0.048 (0.037)	0.056* (0.031)	–0.084 (0.056)
School controls	Yes	Yes	Yes	Yes
School fixed effects	Yes	Yes	Yes	Yes
District-year fixed effects	Yes	Yes	Yes	Yes
Unique schools	943	943	943	942
School-years	4,693	4,693	4,693	4,658

Notes: Asterisks denote significance at the 1% (***), 5% (**), and 10% (*) levels. Data are weighted by size of kindergarten entering class, fall 1998. Robust standard errors are clustered at school level.

educated mothers (Column 1); no corresponding positive trends emerge for children of less educated mothers (Column 2). Likewise, there is evidence for a relative increase in enrollment of non-subsidized lunch-using students (over 185 percent of the poverty line) in the post-policy period for A schools (Column 3), while there is no evidence of similar relative post-policy changes for A schools in terms of enrollment for students who use subsidized lunch.[16]

Overall, these results suggest that parents respond to school grades by enrolling their children in higher-graded schools, and that these responses, especially to A grades, are most pronounced among more affluent, educated parents. These findings on income and education are consistent with the existing literature regarding parental SES and school and teacher choice. Hastings, Kane, and Staiger (2005) find that high-SES families are more responsive to academic quality when exercising school choice, while lower-SES families balance academic considerations with preferences for also attending racially diverse schools, and Jacob and Lefgren (2007) find that low income and minority parents are less likely to actively select a teacher.

C. Does the Performance of Neighbors Affect the School Choice Response to Accountability?

While parents statewide may be expected to respond to new information about the quality of schooling, acting on this information will likely be easier or more urgent for some parents than for others. Specifically, parents located in areas with a greater number of schools nearby may find it easier to respond to the new information because having multiple options nearby may lower the cost associated with choosing a non-zoned school. In addition, parents located in areas where nearby alternatives are of greatly varying quality may find it more urgent to respond to the new information; an A school whose nearest alternatives are predominantly A and B schools may see a smaller influx of new students than A schools located among many C schools. Again, we limit our analyses to looking at the effect of receiving a grade of A.

We look at several measures of the accessibility and quality of nearby schools as potential moderators of the effect of receiving different grade shocks. Our first measure is a simple indicator for whether there is another elementary school within three miles. The rationale behind this measure is that if there are not schools nearby, students will have limited options to transfer. Our second measure is an indicator for whether any school within a three mile radius receives a C or worse. If school grades influence enrollment decisions, the quality of neighboring schools should matter, and we posit that parents will be more likely to seek alternative public schools if their local schools

[16] Although we do not present these results in detail here, the patterns are similar but somewhat less pronounced for B schools.

are receiving grades that parents consider “bad.” A schools with poorer-performing nearby schools should therefore see greater enrollment influxes than A schools with higher-performing neighbors.[17]

Our third measure is an indicator for whether a school’s district received a public school choice incentive grant in 1999. Florida encouraged districts to develop and offer within-district transfer (open enrollment) options in the late 1990s; as part of this effort, the state offered public school choice incentive grants to help districts manage the costs associated with implementing open enrollment plans (Florida House of Representatives, 2001). We hypothesize that parents should be more easily able to respond to school grading information in districts that have a formal process to facilitate open enrollment of students in schools to which they are not zoned. In districts that adhere more rigidly to school zoning as the sole mechanism for school assignment, parents would have to change residences if they wanted to change their children’s school assignment in response to a low school grade. Table 6 describes the share of A, B, and C schools that fall in each of these categories.

The pattern of results (Table 7) suggests that effects of receiving A grades on the advantage index measure are more strongly concentrated where students at neighboring schools may find transfer more desirable or more convenient. Coefficients on the $A \times Year$ variables for the post-policy years are larger and more consistently significant for schools that had other elementary schools within three miles (Column 2) than for schools that had no neighbors in that radius (Column 1). We conducted a joint test of the equality of the $A \times Year$ coefficients for the two groups in the post-policy years; the joint test narrowly misses marginal significance at conventional levels ($p < 0.12$).[18] Findings are quite similar when considering whether an A school had a C or worse-rated school within three miles (Columns 3 and 4). On the other hand, it appears that whether or not a school is in a district with open enrollment, the response to an A grade is quite similar (Columns 5 and 6). In summary, the general pattern of results indicate that the effects of the introduction of school grades on student composition are somewhat more strongly concentrated in the schools positioned to see the greatest parent responses from the new grading system.

[17] In robustness checks, we explore the sensitivity to the use of a 5 mile radius. The results are substantively similar, with no clear patterns for schools that have no competitors/no C-or-lower competitors in 5 miles, and significant positive $A \times Year$ coefficients starting in 1999. We prefer the 3-mile radius because it provides more variability: Relatively few schools (<12% of A–C schools) have no public alternatives within the larger 5-mile radius. We have also characterized school districts based on the distribution of school grades using a Herfindahl index. It does not appear that there is much of a difference in responses to an A grade based on whether a district has a high degree of similarity in school grades or a low degree of similarity in grades across schools. That said, it is somewhat difficult to interpret this Herfindahl measure, because the Herfindahl index would be the same if a school district had, say, one A school and ten C schools, versus one C school and ten A schools, but the interpretation of the value of an A might be much different.

[18] Specifically, we run fully-interacted models with one indicator from each dichotomous pairing and test the $A \times Year$ moderator interactions.

Table 6

Landscape of Nearby Schools as of Fall 1999, by 1999 School Grade

	1 A Mean (Standard Deviation)	2 B Mean (Standard Deviation)	3 C Mean (Standard Deviation)
Any other schools in 3 miles (%)	70.435 (45.854)	70.106 (45.908)	72.745 (44.563)
Any schools C or lower in 3 miles (%)	48.968 (50.231)	49.840 (50.140)	66.478 (47.245)
District had public choice grant: 1999 (%)	57.851 (49.598)	54.430 (49.934)	49.589 (50.037)
Unique schools	114	192	639

Notes: 1998 characteristics are based on 1999 school grades. Data are weighted by 1998 number of students. Average Grade Point Average (GPA) excludes referent school and 2.5 threshold is the median GPA of A schools' neighbors.

D. Sensitivity and Robustness Checks

Because previous versions of school accountability in Florida relied on decisions made at the school district level and because of the considerable variation across school districts in their degree of school choice available, one may be concerned that our results may be driven by results in some subset of school districts, which in Florida are coterminous with counties. We re-run our main advantage index models excluding each county one-by-one, and find largely similar results. While the magnitude of the coefficients changes somewhat as different counties are excluded, the pattern of results — with A schools seeing significant increases in mean advantage in the post-policy years — remains.

Another concern is that the use of a two-thirds cohort (those born January–August 1992, but not their kindergarten classmates born September–December 1991 or any red-shirted or retained classmates) in 1997 may be affecting our findings on pre-policy trends. In particular, this may be a concern to the extent that red-shirted children have different socioeconomic characteristics than non-red-shirted children (Bassok and Reardon, 2013). We therefore re-specify our advantage index and enrollment dependent variables to check whether the pre-policy trends from the main results are reliable. We re-estimate the mix of students in each class and enrollment numbers, using only non-redshirted students born in the first nine months of the year (i.e., those who are comparable to the students represented in our partial cohort entering kindergarten in 1997). Our results (available on request) suggest that our main findings are not masking

Table 7

Effect on Advantage Index of A Receipt for Schools, by Neighbor Quality and District Open Enrollment, Full Sample

	(1) None within 3	(2) Any within 3	(3) None C or Lower within 3	(4) Any C or Lower within 3	(5) No Open Enroll 1999	(6) Open Enroll Grant 1999
	b/se	b/se	b/se	b/se	b/se	b/se
1999 Grade A × 1997	–0.014	–0.003	–0.010	–0.003	–0.016	0.005
	(0.014)	(0.007)	(0.010)	(0.009)	(0.010)	(0.008)
1999 Grade A × 1999	0.009	0.020***	0.010	0.018**	0.013*	0.013*
	(0.011)	(0.006)	(0.008)	(0.008)	(0.007)	(0.007)
1999 Grade A × 2000	0.018	0.027***	0.021**	0.026***	0.020**	0.029***
	(0.013)	(0.007)	(0.009)	(0.008)	(0.009)	(0.007)
1999 Grade A × 2001	–0.000	0.032***	0.002	0.034***	0.020	0.024***
	(0.016)	(0.008)	(0.012)	(0.010)	(0.013)	(0.008)
School controls	Yes	Yes	Yes	Yes	Yes	Yes
School fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
District-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Unique schools	261	636	357	540	474	471
School-years	1,299	3,167	1,776	2,690	2,358	2,346
1997 differential p-value	0.620		0.602		0.092	
1999–2001 joint differential p-value	0.120		0.176		0.844	

Notes: Asterisks denote significance at the 1% (***), 5% (**), and 10% (*) levels. Data are weighted by size of kindergarten entering class, fall 1998. Group outcome means and standard deviations are calculated for group that received grade labeled in column header, in fall of 1998 (weighted by fall 1998 kindergarten student counts). "None within 3" is shorthand for "no other schools within 3 miles." "None C or lower within 3" is shorthand for "no other schools graded C or lower within 3 miles."

any pre-policy trends driven by the change in sample; the pattern of our results is very similar when we use this sample, although consistent with the smaller sample size, the results are less often statistically significant.

As a final robustness check, we explored the sensitivity of results to the exclusion of students who are observed in public-school pre-kindergarten programs run through the K-12 schools, running models like those tested in Table 3. If families are less likely to respond to new information about schools when they have pre-existing ties to the schools, we might see larger effects among this group of students. We find that the pattern of results is the same as for the main sample, as there is no evidence of greater response among this group.

IV. CONCLUSION

One purpose of school accountability systems is to provide information to families that have limited information about the quality of local schools. As such, they are often intended to enhance equity in education. Our results indicate that the families most prepared to act upon this new information may be the most advantaged, and that school accountability could have the unintended consequence of enhancing stratification rather than reducing it. While there exists plentiful evidence that school accountability systems have the potential to improve school quality (Figlio and Ladd, 2008; Figlio and Loeb, 2011), this study provides a cautionary note that has implications for the optimal design of school choice policies in an era of school accountability. Given that rankings implied by school accountability systems are often very weakly correlated with one another (Figlio et al., 2014), the frequently arbitrary decisions that a state makes can have lasting consequences for students and families.

ACKNOWLEDGEMENTS

We are grateful to Eric Brunner, Douglas Lee Lauen, and Sean Reardon, seminar participants at Stanford University, and participants at the 2014 Subnational Government Competition Conference at The University of Tennessee, Knoxville Center for Business & Economic Research and the 2013 Annual meeting of the Association for Education Finance and Policy for helpful comments and suggestions. We are also grateful to the Florida Department of Health and Florida Department of Education for providing the data necessary to carry out this study, and to the Institute for Education Sciences, Bill and Melinda Gates Foundation, and Smith Richardson Foundation for financial support. All errors are our own.

DISCLOSURES

Data were provided by the Florida Department of Education and the Florida Department of Health. Neither author has any financial interests that create conflicts of interest with respect to this paper.

REFERENCES

Bassok, Daphne, and Sean Reardon, 2013. "'Academic Redshirting' in Kindergarten: Prevalence, Patterns, and Implications." *Education Evaluation and Policy Analysis* 35 (3), 238–297.

Beales, Janet R., and Maureen Wahl, 1995. "Given the Choice: A Study of the PAVE Program and School Choice in Milwaukee." Policy Study 183. Reason Foundation, Los Angeles, CA.

Bifulco, Robert, Helen F. Ladd, and Stephen L. Ross, 2009. "The Effects of Public School Choice on Those Left Behind: Evidence from Durham, North Carolina." *Peabody Journal of Education* 84 (2), 130–149.

Billings, Stephen, Eric J. Brunner, and Stephen L. Ross, 2014. "The Housing and Educational Consequences of the School Choice Provisions of NCLB: Evidence from Charlotte, NC." Manuscript. University of North Carolina, Charlotte, NC, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2534040.

Burke, Mary A., and Tim R. Sass, 2013. "Classroom Peer Effects and Student Achievement." *Journal of Labor Economics* 31 (1), 51–82.

Carrell, Scott E., and Mark L. Hoekstra, 2010. "Externalities in the Classroom: How Children Exposed to Domestic Violence Affect Everyone's Kids." *American Economic Journal: Applied Economics* 2 (1), 211–228.

Chakrabarti, Rajashri, 2013. "Do Vouchers Lead to Sorting Under Random Private School Selection? Evidence from the Milwaukee Voucher Program." *Economics of Education Review* (June) 34, 191–218.

Chiang, Hanley, 2009. "How Accountability Pressure on Failing Schools Affects Student Achievement." *Journal of Public Economics* 93 (9–10), 1045–1057.

Chion-Kenney, Linda, 1998. "Schools Scrutinize Initial Test Results." *St. Petersburg Times* May 15, 4B.

Coleman, James S., 1968. "The Concept of Equality of Educational Opportunity." *Harvard Educational Review* 38 (1), 7–22.

Consortium for Policy Research in Education, 2000. "State and District Accountability Systems and the Fifty States: 1999–2000." Consortium for Policy Research in Education, Philadelphia, PA, http://www.cpre.org/images/stories/cpre_pdfs/fl.pdf.

Cowen, Joshua M., and Marcus A. Winters, 2013. "Choosing Charters: Who Leaves Public School as an Alternative Sector Expands?" *Journal of Education Finance* 38 (3), 210–227.

Epple, Dennis, and Richard E. Romano, 2011. "Peer Effects in Education: A Survey of the Theory and Evidence." In Benhabib, Jess, Alberto Bisin, and Matthew O. Jackson (eds.), *Handbook of Social Economics*, *Volume 1B,* 1053–1163. Elsevier, Amsterdam, The Netherlands.

Figlio, David N., 2007. "Boys Named Sue: Disruptive Children and Their Peers." *Education Finance and Policy* 2 (4), 376–394.

Figlio, David N., Jonathan Guryan, Krzysztof Karbownik, and Jeffrey Roth, 2014. "The Effects of Poor Neonatal Health on Children's Cognitive Development." *American Economic Review* 104 (12), 3921–3955.

Figlio, David N., Cassandra M. D. Hart, and Molly Metzger, 2010. "Who Uses a Means-Tested Scholarship and What Do They Choose?" *Economics of Education Review* 29 (2), 301–317.

Figlio, David N., and Lawrence W. Kenny, 2009. "Public Sector Performance Measurement and Stakeholder Support." *Journal of Public Economics* 93 (9–10), 1069–1077.

Figlio, David N., and Helen F. Ladd, 2008. "School Accountability and Student Achievement." In Ladd, Helen F., and Edward B. Fiske (eds.), *Handbook of Research in Education Finance and Policy*, 166–182. Routledge Press, London, UK.

Figlio, David N., and Susanna Loeb, 2011. "School Accountability." In Hanushek, Eric, Stephen J. Machin, and Ludger Woessmann (eds.), *Handbooks in Economics, Volume 3,* 383–421. North-Holland, Amsterdam, The Netherlands.

Figlio, David N., and Maurice E. Lucas, 2004. "What's in a Grade? School Report Cards and the Housing Market." *American Economic Review* 94 (3), 591–604.

Figlio, David N., and Cecilia E. Rouse, 2006. "Do Accountability and Voucher Threats Improve Low-performing Schools?" *Journal of Public Economics* 90 (1), 239–255.

Florida Department of Education, 1999. *School Accountability Report Guide June 1999.* Florida Department of Education, Tallahasse, FL, http://schoolgrades.fldoe.org/reports/9899/guide99.asp.

Florida Department of Education, 2013. "Florida School Accountability Reports." Florida Department of Education, Tallahassee, FL, http://schoolgrades.fldoe.org/.

Florida House of Representatives Council on Lifelong Learning, 2001. "Public School Choice Fact Sheet." Florida House of Representatives, Tallahassee, FL, http://www.leg.state.fl.us/data/publications/2002/house/reports/EdFactSheets/fact%20sheets/Public%20School%20Choice.pdf.

Greene, Jay P., William G. Howell, and Paul E. Peterson, 1997. *Lessons from the Cleveland Scholarship Program.* Paper presented at the Annual Meeting of the Association of Public Policy and Management, November, sponsored by the Taubman Center on State and Local Government and the Center for American Political Studies, Washington, DC.

Hastings, Justine S., Thomas J. Kane, and Douglas O. Staiger, 2005. "Heterogeneous Preferences and the Efficacy of Public School Choice." NBER Working Paper 11805. National Bureau of Economic Research, Cambridge, MA.

Hastings, Justine S., and Jeffrey M. Weinstein, 2008. "Information, School Choice, and Academic Achievement: Evidence from Two Experiments." *Quarterly Journal of Economics* 123 (4), 1373–1414.

Hoxby, Caroline M., and Gretchen Weingarth, 2005. "Taking Race out of the Equation: School Reassignment and the Structure of Peer Effects." Manuscript. Harvard University Department of Economics, Cambridge, MA.

Imberman, Scott A., Adriana D. Kugler, and Bruce Sacerdote, 2012. "Katrina's Children: Evidence on the Structure of Peer Effects from Hurricane Evacuees." *American Economic Review* 102 (5), 2048–2082.

Jacob, Brian A., and Lars Lefgren, 2007. "What Do Parents Value in Education? An Empirical Investigation of Parents' Revealed Preferences for Teachers." *Quarterly Journal of Economics* 122 (4), 1603–1637.

Lavy, Victor, M. Daniele Paserman, and Analia Schlosser, 2012. "Inside the Black Box of Ability Peer Effects: Evidence from Variation in the Proportion of Low Achievers in the Classroom." *Economic Journal* 122 (559), 208–237.

Lewis, Jeffrey B., and Drew A. Linzer, 2005. "Estimating Regression Models in Which the Dependent Variable is Based on Estimates." *Political Analysis* 13 (4), 345–364.

Reardon, Sean F., 2011. "The Widening Academic Achievement Gap Between the Rich and the Poor: New Evidence and Possible Explanations." In Duncan, Greg G., and Richard J. Murnane (eds.), *Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances,* 91–116. Russell Sage Foundation Press, New York, NY.

Rothstein, Jesse M., 2006. "Good Principals or Good Peers? Parental Valuation of School Characteristics, Tiebout Equilibrium, and the Incentive Effects of Competition among Jurisdictions." *American Economic Review* 96 (4), 1333–1350.

Rouse, Cecilia Elena, Jane Hannaway, Dan Goldhaber, and David N. Figlio, 2013. "Feeling the Florida Heat? How Low-performing Schools Respond to Voucher and Accountability Pressure." *American Economic Journal: Economic Policy* 5 (2), 251–281.

Schneider, Mark, and Jack Buckley, 2002. "What Do Parents Want from Schools? Evidence from the Internet." *Educational Evaluation and Policy Analysis* 24 (2), 133–144.

Sojourner, Aaron, 2013. "Identification of Peer Effects with Missing Peer Data: Evidence from Project STAR." *Economic Journal* 123 (569), 574–605.

West, Martin R., and Paul E. Peterson, 2006. "The Efficacy of Choice Threats within School Accountability Systems: Results from Legislatively Induced Experiments." *Economic Journal* 116 (510), C46–C62.

Witte, John, 2001. *The Market Approach to Education: An Analysis of America's First Voucher Program.* Princeton University Press, Princeton, NJ.

Zimmer, Ron, Brian Gill, Kevin Booker, Stephane Lovertu, Tim R. Sass, and John Witte, 2009. *Charter Schools in Eight States: Effects on Achievement, Attainment, Integration, and Competition.* RAND Corporation, Santa Monica, CA.